IZA DP No. 1692

Don't Give Up On Me Baby: Spousal Correlation in Smoking Behaviour

Andrew E. Clark
Fabrice Etilé

July 2005

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Don’t Give Up On Me Baby: Spousal Correlation in Smoking Behaviour

Andrew E. Clark
CNRS, PSE and IZA Bonn

Fabrice Etilé
INRA-CORELA

Discussion Paper No. 1692
July 2005

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 1692
July 2005

ABSTRACT

Don't Give Up On Me Baby: Spousal Correlation in Smoking Behaviour*

We use nine waves of BHPS data to examine interactions between spouses in terms of a behaviour with important health repercussions: cigarette smoking. Correlation between partners' behaviours may be due to correlated effects, as a consequence of matching or information revealed by others' behaviours, or to endogenous effects generated by bargaining within marriage. A simple bivariate probit reveals a positive correlation between own current smoking and partner's past smoking, which is consistent with endogenous effects. However, after controlling for individual effects, we find that own current smoking and partner's past smoking are statistically independent: all of the correlation in smoking status works through the correlation in individual effects. As such the correlation in the raw smoking data is consistent with positive assortative matching in marriage over smoking, rather than bargaining within the couple or social learning.

JEL Classification: C33, D83, I12, I18

Keywords: smoking, matching, bargaining, learning, health

Corresponding author:

Fabrice Etilé
INRA-CORELA
65 Boulevard de Brandebourg
94205 Ivry-sur-Seine cedex
France
Email: etile@ivry.inra.fr

* We thank Andrew Jones, Brian Krauth, François Legendre, Maarten Lindeboom, Nigel Rice, Frank Windmeijer, and seminar participants at INRA-CORELA, the 20th Journées de Micro-économie Appliquée (Montpellier), the York Seminar in Health Econometrics, and the 12th European Workshop on Econometrics and Health Economics (Menorca). We are particularly grateful to David Madden and two anonymous referees for very detailed comments. Paris-Jourdan Sciences Economiques (PSE) is a Joint Research Unit CNRS-EHESS-ENPC-ENS. The BHPS data were made available through the ESRC Data Archive. The data were originally collected by the ESRC Research Centre on Micro-social Change at the University of Essex. Neither the original collectors of the data nor the Archive bear any responsibility for the analyses or interpretations presented here. This research was supported by grants from the Institut National de la Santé et de la Recherche Médicale (INSERM) and the Mission Interministerielle à la Lutte contre la Drogue et la Toxicomanie (MILDT).

Don't Give Up on Me Baby: Spousal Correlation in Smoking Behaviour

Andrew Clark and Fabrice Etilé

1 Introduction

It is perhaps a commonplace to say that the behaviours of individuals within a household tend to be similar. For instance, two recent papers, Farrell and Shields (2002) and Leonard and Mudar (2003), have uncovered empirical evidence of spousal correlation in sporting activity and drinking respectively. Along the same lines, Latkin *et al*. (1995) find evidence of spousal correlation in injecting habits amongst drug users, while Kan and Heath (2003) use British panel data to investigate similarities in husbands' and wives' political preferences. Fernández *et al.* (2005) show a positive correlation between spouse's years of schooling across 34 countries, while Rose (2004) finds evidence of increasing positive assortative matching with respect to education in American data. Last, Jenkins and Osberg (2005) consider the joint production of leisure in couples.

In this paper, we examine interactions between spouses[1] in terms of an observable behaviour with important repercussions on health: cigarette smoking. In long-run panel data we find, as expected, a strong correlation between husbands' and wives' smoking. Following Manski's (1995) approach to observed social interactions, this raw correlation can reflect three separate phenomena. First, individuals may tend to marry those who share the same tastes and characteristics. This is analogous to the concept of positive assortative matching in Becker (1974), and will induce correlated effects in Manski's terminology.

Second, correlated effects may also arise because spouses share the same environment. In particular, they face the same prices and local environment, and largely receive the same informations. In this context, Clark and Étilé (2002) consider interactions *via* learning about health risks, whereby the health developments of other household smokers reveal information about one's own risk from smoking.

Last, correlation in partners' smoking may reflect the household's decision-making process. The idea here is that the presence of some shared marital output, which is affected by smoking, may lead couples to interact via their smoking status in cooperative or non-cooperative bargaining. This can be likened to the endogenous effects in Manski (1995).

---

[1] We use the terms "spouse" and "marriage" loosely here, as we consider both legally married couples and those who live together.

This paper uses nine waves of British Household Panel Survey (BHPS) data to look at the correlation in smoking between partners. We consider both smoking participation and, using the panel aspect of the data, the quit decision. We estimate a number of different specifications to try to distinguish between correlated and endogenous effects. Assortative matching on lifestyle preferences will be picked up by correlated individual effects in the male and female smoking equations.

Under household decision-making, we expect partners' behaviours to be correlated, even after individual random effects have been introduced (these latter will only pick up the time-invariant correlation between spouses' preferences). This will be our key test of endogenous versus correlated effects. We will also test directly for learning by including partner's past health developments.

Our results show that almost all of the correlation in partners' smoking comes from the positive correlation between their individual effects. This suggests that smoking, and perhaps more generally health or lifestyle, is one of the domains over which sorting takes place in the marriage market. Further, the positive correlation shows that lifestyles are complements in the household production function: there is positive assortative matching over smoking.

The paper is organised as follows. We first run through some ways in which to think of correlation between partners' behaviour: matching, social learning and household bargaining. Section 3 then presents the data, and Section 4 the econometric approach. Our main findings are reported in Section 5. We show that matching explains the spousal correlation in smoking behaviour. Section 6 proposes further empirical results regarding learning and the presence of children, and Section 7 concludes.

2 Theory

We briefly present here a number of arguments which may account for correlation between partners' smoking behaviour. These are not specific to cigarette consumption, and may apply more generally to other behaviours (see also Wilson, 2002). They fall into three different groups: assortative matching in the marriage market; social learning about smoking's health risks from the observation of one's partner; and interactions due to bargaining within marriage. We then outline the empirical tests which we will appeal to in our estimations to distinguish between these separate interpretations.

2.1 Matching in the Marriage Market

The first theoretical consideration refers to the process of matching on the marriage market. A widely-cited article by Becker (1974) considers the gains from marriage accruing to two rational individuals. Assuming transferable utility, it is possible to define an output measure characterising the gains from matches on the marriage market (Bergstrom, 1997). We consider that smoking, along with other goods and household members' time inputs, contributes to the production of marital output. One major implication of Becker's theory of marriage is that complementarity of partners' traits in the marital production function implies positive assortative matching. Smoking may be considered as one of the traits that determine marriage assignments, or, more generally, as an easily observed signal of (less-easily observed) general preferences over other activities and goods, such as parties, concerts, healthy foods and sport. Contoyannis and Jones (2004) show that a number of such lifestyle variables are correlated between themselves. It seems likely that these lifestyle variables will be complements in the marital production function, in the sense that partners enjoy sharing these activities. As such, we expect positive assortative matching with respect to lifestyle preferences, including smoking.

A related point concerns matching with respect to life expectancy, which latter is reduced by smoking. Risk-aversion to time spent alone in widowhood will reinforce preferences for partners whose life expectancy coincides with ones own.

Matching may also cover observable traits like earnings. Here, a number of studies have found a wage penalty for smokers (see Levine *et al*., 1997, and Van Ours, 2004) with the impact of health on labour market outcomes being more pronounced for men than for women (Currie and Madrian, 1999). Becker's model predicts negative assortative matching for earnings, as a result of human capital specialization in the market or non-market sectors. Hence, smokers may be *ceteris paribus* more attractive for non-smokers on the marriage market, via specialisation in the non-market and market sectors respectively. This is one potential explanation for mixed matches between smokers and non-smokers.

Matching on the marriage market thus corresponds to correlated effects in partners' behaviours. The implication for empirical estimation is that partners' smoking will be correlated due to similarities in unobservable individual traits. We control this by including correlated individual random effects in both male and female smoking equations. Our *a priori* is that the positive assortative matching from lifestyles will outweigh the negative assortative matching from wages, so that the individual effects will be positively correlated.

Alternatively, correlated effects may result from similarities in the partners' information sets, as described below.

2.2 Social Learning

The argument here applies to a world in which there is uncertainty about the risk from smoking; as such new information regarding this risk may be received by individuals, in particular through the observation of others who smoke. Partners will likely learn from each other, so that their information will be shared, and their risk assessments will be correlated. This leads naturally to a correlation in observed behaviour.

Information is difficult to measure, as indeed are risk assessments. Our empirical approach will control for correlated information by allowing for correlations in contemporaneous unobservable shocks: specifically, male and female smoking equations are written as a bivariate probit where the error terms are not necessarily independent. In Section 6, we will also look for direct evidence of correlated information effects using the approach taken by Clark and Etilé (2002), wherein health changes for one partner who smokes may affect the perceived dangers from smoking for both.

2.3 Household Decision-Making

The last interpretation of the correlation between partners' behaviour relies on the ongoing decision-making process within the household. Following Manser and Brown (1980) and McElroy and Horney (1981), spouses may have conflicting preferences over smoking. Smoking's payoffs to the individual depend on both the private pleasure from smoking and the effect of both partners' smoking on the production of marital output. In this paper, we exclude *a priori* that conflicts over smoking decisions may induce divorce. Hence, the individual threat points are the highest payoffs that each spouse can achieve by acting strategically within marriage (*cf.* Bolin *et al.*, 2002, for a theory of partners' strategic decision-making over health). Of course, spouses can also cooperate, and achieve a Pareto-efficient outcome through, for instance, a Nash bargaining process that ensures utility greater than these threat points (*cf.* Bolin *et al.*, 2001, for a theory of Nash bargaining over household health investments).

To illustrate this argument, consider a basic set-up such as the battle of the sexes, where partners' payoffs are given by the following matrix and are common knowledge. The first entry in each cell is the male's payoff and the second entry is the female's payoff.

		Male	
		Smokes	*Does not smoke*
Female	*Smokes*	B^M-2p-H, B^F-2p-H	-p-H, B^F-p-H
	Does not smoke	B^M-p-H, -p-H	0,0

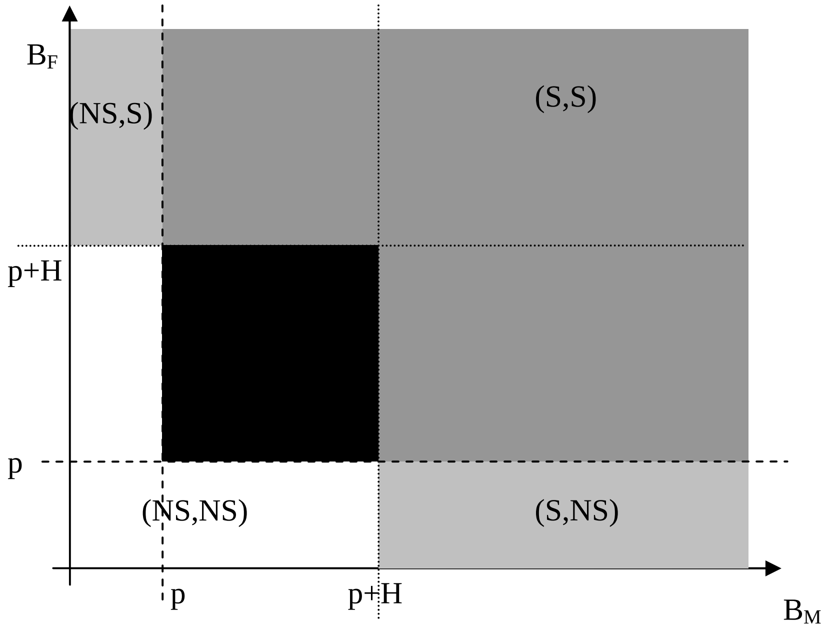
Here, B^M and B^F are respectively the male's and female's private returns from smoking. The financial cost of each partner's smoking, *p*, affects both spouses. Last, *H* is the health impact of smoking, which is shown in the payoff matrix as being shared between spouses. The idea here, which is simple for presentational purposes, is that both spouses need to be in good health for some marital outputs to be produced. If one of the two smokes, then this output is lost, with a subsequent health cost for each partner of *H*.[2]

Suppose first that spouses act strategically. The solution, as shown in Figure 1, depends on the distribution of the private returns B^M and B^F. For example, a high value of B^F (greater than $p+H$) coupled with a low B^M (under *p*) produces (NS, S), meaning that the male does not smoke but his partner does. Outside of the black area, the non-cooperative solution is Pareto dominant, so that there is no benefit from cooperating.

When private payoffs are in the black area, there is indeterminacy about the optimal solution since both (smoking, smoking) and (not smoking, not smoking) are pure strategy Nash equilibria. To solve this indeterminacy, we may assume that the mixed-strategy equilibrium defines appropriate threat-points. The black area thus delineates the payoffs for which cooperative bargaining may bring about the Pareto-dominant solution (here, the not smoking equilibrium).

[2] The health effect, *H*, is presented here as being non-separable between spouses (for instance the value of time spent together in good health). We can also imagine strictly private health costs. For the bargaining explanation to be relevant, one of the financial (*p*) or health (*H*) costs needs to be shared.

Figure 1: Private payoffs and optimal solution to the strategic game.



Household decision-making has major implications for empirical modelling of household smoking. The smoking status of one spouse at *t* will depend on the two private returns, which themselves, *via* addiction, depend on both spouses' smoking statuses at time *t-1*.

Any household decision process over smoking for which the threat-points do not involve re-marriage can be distinguished from matching, because the outcome is susceptible to change over time; further, any change in own smoking should be systematically related to changes in both partners' past smoking. The empirical interpretation is that respondent smoking at *t* depends on partner smoking at *t-1*, even after controlling for individual random effects. This effect is expected to be positive, since there is complementarity both between partners' health statuses in household production, and between past and current smoking statuses in the personal return from smoking.[3] If this relationship is not found in the data, we will conclude that household bargaining does not explain the correlation in spousal smoking.

Any such relationship can also be interpreted more broadly in terms of household bargaining over life expectancy, and not only smoking. Under uncertainty, one important benefit from marriage is risk-sharing (Weiss, 1997). In this respect, aversion to spending time on one's own in widowhood may lead spouses to under-invest in health. If one's partner's

---

[3] This is an assumption of the rational addiction model (Becker and Murphy, 1988). If U is the (separable) private utility of smoking, then $B(S_{t-1})=U(1,S_{t-1})-U(0,S_{t-1})$ is the (net) private return from smoking at time *t*, where S_{t-1} is a binary indicator for past smoking. Complementarity of past and present smoking means here that $B(1)=U(1,1)-U(0,1)$ is higher than $B(0)=U(1,0)-U(0,0)$. Hence, due to addiction, the household's position in Figure 1 drifts to the North or to the East if one of the partners smokes.

health investments are not perfectly observable, smoking may be interpreted as a signal of the partner's commitment to healthy behaviours. In this game, spouses may tend to end up, as in Figure 1, in homogamous equilibria where both smoke or, especially when there is cooperation, no-one smokes. As past health changes are a good proxy for private health investments, we will include them in some of our regressions to control for this signalling effect of smoking.

Last, the solution in Figure 1 depends on the relative size of the private and public payoffs; life-cycle events can change smoking decisions by altering these payoffs. In particular, the presence of children may raise the value of time spent together (and thus increase *H*). This shift in the relative size of the public and private payoffs makes it more likely that the Non-Smoker/Non-Smoker solution Pareto-dominates. However, it is possible that gender asymmetries may yield lower correlations in spousal smoking: mother's smoking has a greater impact on children's health than father's smoking. In Section 6, we will explore the impact of children on smoking outcomes.

2.4 Empirical Implications

Sections 2.1 through 2.3 presented three different explanations of spousal correlation in smoking. These provide us with arguments for the specification of an appropriate household smoking model in Section 4. Our two dependent variables, partners' smoking statuses at time *t*, are binary. Matching implies correlation in individual traits, which leads us to a model with individual random effects that are correlated between partners. Further, social learning suggests that the error terms in male and female smoking equations may be correlated. Together, these suggest a bivariate probit with random effects.

Regarding the right-hand side variables, household decision-making implies that both own and partner's lagged consumption will be important. Last, both social learning and household decision-making (in its health investment incarnation) argue for the use of partner's health developments as an explanatory variable, although for different reasons.

These arguments yield three empirical predictions:

Prediction 1: If there is positive assortative matching over smoking, partners' estimated individual random effects will be positively correlated.[4]

---

[4] That is, we expect the positive assortative matching on lifestyle preferences to be stronger than the negative assortative matching on earnings.

Prediction 2: Household decision-making suggests that respondent's current smoking status should be positively correlated with partner's lagged smoking, once individual random effects are controlled for.

Prediction 3: Partner's health developments may affect respondent smoking. In the case of social learning, the correlation will be positive if the partner smokes and zero if he does not (in the latter case any change in partner's health cannot be attributed to cigarettes). In the health investment game, the correlation will be negative (the individual smokes more to bring her life expectancy into line with her partner's).

In the following sections, we will try to evaluate these three predictions using long-run British panel data.

3 Data

3.1 The British Household Panel Survey

The British Household Panel Survey is an annual panel of roughly 10 000 individuals in around 5 000 different households in Great Britain. We use the first nine waves (1991-1999) of data. All adults in the household are interviewed separately with respect to their socio-demographic characteristics, income, employment, and health. Further details of this survey are available at the following address: http://www.iser.essex.ac.uk/bhps.

3.2 Smoking in the BHPS

We first consider all individuals observed over at least two consecutive periods, to have an idea of how smoking and marital status interact. This initial unbalanced subsample (Sample 1) includes 63530 observations (12467 individuals) of which:

- 21172 are not in a couple at both $t-1$ and t.
- 2814 change from not being in a couple to being in a couple (married or living together), or vice versa, between $t-1$ and t.[5]
- 39544 remain with their partner between $t-1$ and t.[6]

[5] There will likely be a number of couples who do not live together at $t-1$, and who subsequently live in the same house (either cohabiting or married) at t. As we cannot identify the first status, these individuals will be counted in this second category.

[6] A couple who live together at $t-1$, and are married at t, will appear in this third category.

Smoking participation by marital status and sex is summarised in Table 1 (where "single" refers also to those who did not remain with the same partner between *t-1* and *t*).

Table 1: Smoking participation by marital status and sex (Sample 1).

	No. Observations (% Smokers)
Men in couples	19890 (25.8%)
Single Men	9802 (33.1%)
All Men	**29692 (28.3%)**
Women in couples	19654 (24.3%)
Single Women	14184 (30.0%)
All Women	**33838 (26.7%)**
All	***63530 (27.4%)***

While two-thirds of men interviewed in this sample are in couples, this is true for under sixty per cent of women. There is a difference of 7.3% for men in the smoking participation of singles versus couples; the analogous figure for women is 5.7% per cent. It is possible that this reflects an indirect widowhood effect. Men in couples are about 9.5 years older than single men, while single women are only three years older than women in couples. Widowhood explains these differences (32% of single women are widowed, against only 11% of single men).

For the regression analysis, we consider couples who stay together over all nine waves, and for whom information on both partners (of different gender) is available. As we also require one-period lagged smoking, this leaves us with roughly 10500 observations on couples observed over eight periods in this balanced sub-sample (Sample 2).

Descriptive statistics for samples 1 and 2 are shown in Appendix A. Individuals in stable couples (those in sample 2) are less likely to smoke, are richer, have more children, and are less likely to cohabit. They are also less likely to suffer from unemployment and poor health.

Smoking participation for men and women in Sample 2 is 20.9% and 20.3% respectively (which is four or five percentage points lower than the figures for Sample 1 in Table 1). The crosstabulation of couples' smoking statuses is shown in Table 2 below.

Table 2: Couples' smoking statuses.

		Male	
		Smoker	*Non-Smoker*
Female	*Smoker*	10.7%	9.6%
	Non-Smoker	10.2%	69.5%

The table reveals that there are more mixed couples (one smokes, the other does not) than there are matched smoking couples, suggesting that many factors, in addition to lifestyle preferences, are important in couple formation.

Table 3: Conditional Probabilities.

		Female	
		Smoker	*Non-Smoker*
Male	*Smoker*	52.8%	12.8%
		Male	
		Smoker	*Non-Smoker*
Female	*Smoker*	51.2%	12.1%

We can calculate conditional probabilities to illustrate the correlation in spouses' behaviours. Table 3 presents these figures, conditional on partner's smoking status. The table should be read as follows: given that the woman smokes, the probability that the man smokes is 52.8%; given that the female partner does not smoke, the probability that the male partner smokes is only 12.8%. The conditional probabilities for women tell almost the same story. Nevertheless, the descriptive odds ratio for women is a little higher than that for men (4.23 and 4.12 respectively)[7]: male smoking might be more of a risk factor for women than is female smoking for men. There is obviously a positive correlation between partners' smoking statuses, with a slight gender asymmetry.[8] In the remainder of this paper, we try to investigate this correlation in the light of section 2's theoretical considerations.

3.3 Accounting for selection bias

Although we are interested only in couples, and the threat of divorce is excluded from our theoretical considerations (see Section 2.3), there is likely some selection bias involved in moving from sample 1 to sample 2, as a number of contemporaneous shocks may simultaneously affect the duration of the couple and smoking behaviour. As such, the regression sample will not necessarily reflect the total population of couples. In particular,

[7] The descriptive odds ratio is the ratio of the conditional probability of positive outcome when the conditioning variable is active to the conditional probability of positive outcome when the conditioning variable is inactive. Here: $\frac{\Pr(\text{male smokes} \mid \text{female smokes})}{\Pr(\text{male smokes} \mid \text{female does not smoke})} = \frac{0.528}{0.128} = 4.12$.

[8] This (small) difference in the descriptive odds ratio is consistent with the lower wage penalty for women smokers described above.

there could be an unobserved variable which determines couple stability, and which is also linked to their joint smoking status.

To correct for any selection bias in moving from sample 1 to sample 2, we compute a Mills ratio using a selection variable that equals 1 at period *t* if the individual is observed over the 9 periods and does not separate in period *t*. This marital status selection equation is estimated on sample 1, as shown in Appendix Table B1, as a function of education (3 dummies), labour force status (10 dummies), region and year, and regional unemployment rates by sex and year. These last variables are used to satisfy the exclusion restrictions, which is possible in our theoretical framework since the threat-points are defined by non-cooperative behaviour within marriage, and not by divorce.

4 Econometric Modelling: A Dynamic Bivariate Probit Model

Let $Y_{i,t}$ be a binary indicator for smoking by individual *i* during period *t*, and $X_{i,t}$ a vector of exogenous individual and household covariates. The agent decides to smoke at time *t* ($Y_{i,t}$=1) if the latent variable $Y^*_{i,t}$ is positive. We consider the following bivariate probit specification of the household smoking decision:

$$Y_{i,t} = \begin{cases} 1 & \text{if} \quad Y^*_{i,t} = \alpha_1 Y_{i,t-1} + \beta_1 Y_{-i,t-1} + \gamma_1 X_{i,t} + c_i + \tilde{\varepsilon}_{i,t} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$Y_{-i,t} = \begin{cases} 1 & \text{if} \quad Y^*_{-i,t} = \alpha_2 Y_{-i,t-1} + \beta_2 Y_{i,t-1} + \gamma_2 X_{-i,t} + c_{-i} + \tilde{\varepsilon}_{-i,t} > 0 \\ 0 & \text{otherwise} \end{cases} \qquad \textbf{(1)}$$

Here $Y_{-i,t-1}$ refers to the lagged smoking status of *i*'s partner. The residuals have two components: an individual specific effect c_i and a time-varying random shock $\tilde{\varepsilon}_{i,t}$. In terms of the three arguments presented in section 2, the correlation between c_i and c_{-i} will capture matching, whereas β corresponds to the bargaining effect. Last, $(\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{-i,t})$ are bivariate normally distributed (with variances normalised to 1), which captures time-varying correlated effects.

Index by *h,t* all column vectors whose two components are the variables describing the behaviours or characteristics of each spouse at time *t*, for instance $Y_{h,t}' = (Y_{i,t}, Y_{-i,t})$. We are interested in the identification of the λ's and γ's in the following model:

$$\Pr(Y_{h,t} \mid Y_{h,t-1}, X_{h,t}, c_h) = \Phi_2[(2Y_{i,t}-1)(\lambda_1 Y_{h,t-1} + \gamma_1 X_{i,t} + c_i), \\ (2Y_{-i,t}-1)(\lambda_2 Y_{h,t-1} + \gamma_2 X_{-i,t} + c_{-i}), (2Y_{i,t}-1)(2Y_{-i,t}-1)\rho] \quad \textbf{(2)}$$

where $\lambda_1=(\alpha_1, \beta_1)$, $\lambda_2=(\alpha_2, \beta_2)$, Φ_2 is the standard normal bivariate c.d.f. and ρ is the correlation coefficient between the shocks $\tilde{\varepsilon}_{i,t}$ and $\tilde{\varepsilon}_{-i,t}$.

Assume that the vector of individual effects has a conditional density $f(c_h|Y_{h,0},X_h)$, where X_h is the vector $(X_{h,1},X_{h,2},\ldots,X_{h,T})'$. We can write the log-likelihood of observations $Y_{h,t}$, $t=1$ to T as follows (Wooldridge, 2002*a*, 2002*b*): [9]

$$\Pr(Y_{h,1}, Y_{h,2}, ..., Y_{h,T} \mid Y_{h,0}, X_h) = \int_{\mathrm{supp}(c_h)} \left[\prod_{t=1}^{T} \Pr(Y_{h,t} \mid Y_{h,t-1}, X_{h,t-1}, c_h) \right] f(c_h \mid Y_{h,0}, X_h) dc_h \quad \textbf{(3)}$$

To estimate this model, we will assume that *f* is a bivariate discrete distribution. Hence c_i takes a finite number S_1 of values c_j on the real line. Accordingly, S_2 is the number of support points c_k for the marginal distribution of c_{-i}.

The conditional probability θ_{jk} that $c_h=(c_i,c_{-i})$ is equal to $c_{jk}=(c_j, c_k)$ is modelled as a multinomial logit, with $Z_h=(Y_{h,0}, X_h)$ as regressors (Greene, 2002). Hence, there are vectors of coefficients $\delta_{11}, \delta_{12},\ldots, \delta_{S_1S_2}$ such that:

$$\theta_{jk} = \Pr(c_h = c_{jk}) = \frac{\exp(\delta_{jk} Z)}{\sum_{j=1}^{S_1} \sum_{k=1}^{S_2} \exp(\delta_{jk} Z)} \quad \textbf{(4)}$$

with δ_{11} normalised to 0. The final household likelihood is:

$$\Pr(Y_{h,1}, Y_{h,2}, ..., Y_{h,T} \mid Y_{h,0}, X) = \sum_{j=1}^{S_1} \sum_{k=1}^{S_2} \theta_{jk} \left[\prod_{t=1}^{T} \Pr(Y_{h,t} \mid Y_{h,t-1}, X_{h,t-1}, c_{jk}) \right] \quad \textbf{(5)}$$

The parameters c_{jk}, α, β, δ and γ can then be identified from the data, by maximisation of the log-likelihood.

The model is estimated by the Simulated Annealing EM algorithm (Celeux *et al.*, 1995), which is a stochastic version of the EM algorithm of Dempster *et al.* (1977). The EM algorithm is increasing in lnL but is known to lead often to poor local maxima or saddle points. The SAEM algorithm was designed to overcome this limitation. Following other authors, we compare information criteria such as the BIC, and the AIC for 2, 3 or more points

[9] This conditioning technique is proposed by Wooldridge to deal with the “initial conditions” problem that arises as a consequence of the correlation between $Y_{h,0}$ and the fixed effect in $\Pr(Y_{h,1}| Y_{h,0},X_{h,1},c_h)$.

c_{jk} in order to find out the optimal numbers S_1 and S_2 of points of support (see for instance Deb and Trivedi, 1997 or Wedel *et al.*, 1993).[10] An entropy-based measure is also used as an indicator for the satisfactory distribution of the individual heterogeneity over the support points[11]. The closer this is to 0, the more inaccurate is the classification of observations into distinct homogeneous groups (Jedidi *et al.*, 1997). Last, note that our model may be misspecified if two-period lagged decisions affect current smoking statuses. As we use a finite discrete distribution to model household specific effects, there is no test of omitted variables: standard tests do not account for uncertainty over the optimal numbers of support points, which may vary with the set of regressors. We will however compute LM statistics for omission of $Y_{h,t-2}$, but they should be considered cautiously.

Table 4 below illustrates our search for the optimal number of support points in the discrete distribution of individual heterogeneity. The information criterion with sample penalties (BIC) suggests four (2 x 2) as the "best" number of points, whereas AIC favours nine (3 x 3) support points. However, the entropy is higher with four support points, and with nine support points, two classes turn out to be only sparsely populated (with a mass of about 2%) and several class memberships are not well identified (the variance of the δ_{jk} in equation **(4)** is very high). Hence, we retain four support points.

Table 4: Random-effect Bivariate Probit Model – selection of the number of mass points.

S_1	1		2			3	
S_2	1	2	1	2	3	2	3
Log-likelihood	-2774	-2586	-2573	-2378	-2334	-2334	**-2271**
BIC	-2988	-2868	-2856	**-2823**	-2941	-2941	-3118
AIC	-2817	-2647	-2635	-2474	-2465	-2465	**-2454**
Entropy	100%	93.0%	92.1%	**94.9%**	90.2%	90.8%	80.7%

In Table 5's specification with individual effects (specification **2**), the variances of the coefficients are computed under the assumption that the optimal number of support points is not a random variable. A rigorous computation of the variance-covariance matrix should consider this issue. To our knowledge, this has not been treated in the existing literature, and we follow tradition by using the standard information matrix.

[10] AIC=LnL-p where LnL is the log-likelihood and p the number of parameters in the model. BIC=LnL-LnN*p/2 where N is the number of observations. There is no test of the optimal number of classes as all support points have non-zero mass *i.e.* the conditions of compactness of the parameter space are not met.

Last, in bivariate probits the use of control variables which take different values for the two partners (labour force status, age etc.) allows the robust identification of the correlation coefficient (Keane, 1992). For the sake of comparison, we will also estimate a bivariate probit model without individual random effects (specification **1**).

5 Matching or Bargaining?

Table 5 below reports results from two different specifications. The table has 4 columns. Columns 1 and 2 report benchmark estimates from a bivariate probit specification without fixed effects. The random effects bivariate probit in columns 3 and 4 adds an individual random effect for both partners, with these random effects following a finite discrete distribution.[12]

Table 5. Matching vs. Bargaining.

Specification	**1: Bivariate Probit**		**2: Discrete Random Effect Bivariate Probit**	
	Male	*Female*	*Male*	*Female*
Past participation: $Y_{i,t-1}$	3.110*** (0.050)	3.526*** (0.058)	1.926*** (0.125)	2.070*** (0.181)
Partner's past participation: $Y_{-i,t-1}$	0.307*** (0.053)	0.300*** (0.063)	0.097 (0.422)	0.125 (0.536)
Mills Ratio	0.283** (0.135)	0.024 (0.150)	0.174 (0.220)	-0.209 (0.386)
Partner's Mills ratio	0.248* (0.132)	0.361** (0.180)	0.100 (0.245)	0.280 (0.323)
Age/10	0.304** (0.143)	0.095 (0.147)	0.694*** (0.170)	0.159 (0.275)
$Age^2/100$	-0.038*** (0.014)	-0.014 (0.014)	-0.068*** (0.019)	-0.017 (0.032)
Log (real income)	-0.133*** (0.036)	-0.096*** (0.029)	-0.065 (0.044)	-0.058 (0.036)
Education ≥ A-level	-0.112** (0.051)	-0.105* (0.057)	-0.140*** (0.052)	0.003 (0.072)
Has at least one child at home	0.009 (0.065)	0.143* (0.075)	-0.007 (0.075)	0.132 (0.116)
Newborn child between *t-1* and *t*	-0.026 (0.153)	-0.343 (0.216)	-0.012 (0.372)	-0.538* (0.307)
Living together (ref: married)	0.067 (0.128)	0.174 (0.142)	0.447** (0.183)	0.232 (0.207)

[11] Entropy= $1-\frac{\sum_h\sum_j\sum_k -\varpi_{hjk}\ln(\varpi_{hjk})}{N*\ln(S^*)}$ where ϖ_{hjk} is the probability that $c_h=c_{jk}$ given the information available in the dataset.

[12] As shown in Table 4, the usual information criteria (AIC or BIC) favour the random-effects bivariate probit over the simple bivariate probit in columns 1 and 2. The estimated results from the multinomial logit in equation (4) are shown in Appendix Table B2.

Mother/father at home	-0.679** (0.267)	-0.313*** (0.080)	-0.649 (0.491)	-0.373*** (0.127)
Year = 1993 (ref: 1992)	0.090 (0.097)	-0.019 (0.118)	-0.057 (0.141)	-0.200 (0.279)
Year = 1994 (ref: 1992)	0.183* (0.099)	-0.032 (0.115)	-0.055 (0.209)	-0.392 (0.262)
Year = 1995 (ref: 1992)	0.254*** (0.097)	0.281*** (0.106)	-0.029 (0.128)	-0.085 (0.253)
Year = 1996 (ref: 1992)	0.281*** (0.097)	0.132 (0.109)	-0.001 (0.136)	-0.228 (0.244)
Year = 1997 (ref: 1992)	0.117 (0.097)	0.127 (0.108)	-0.198 (0.159)	-0.253 (0.260)
Year = 1998 (ref: 1992)	0.116 (0.099)	0.182* (0.105)	-0.208 (0.160)	-0.145 (0.331)
Year = 1999 (ref: 1992)	0.154 (0.098)	0.090 (0.112)	-0.178 (0.114)	-0.257 (0.197)
Constant c_{11} (average probability $\bar{\theta}_{11}$ =9.0%)	-1.918*** (0.525)	-1.962*** (0.496)	-1.701** (0.714)	-0.205 (1.064)
Constant c_{12} ($\bar{\theta}_{12}$ =13.3%)	No	No	-1.701** (0.714)	-2.413** (0.986)
Constant c_{21} ($\bar{\theta}_{21}$ =11.9%)	No	No	-3.665*** (0.709)	-0.205 (1.064)
Constant c_{22} ($\bar{\theta}_{22}$ =65.8%)	No	No	-3.665*** (0.709)	-2.413** (0.986)
Rho / Gender correlation in time-varying random errors	0.521*** (0.046)		0.591*** (0.129)	
Controls for initial conditions: $Y_{h,0}$, $X_{h,t}$ (t=1,…,8)	No		Yes	
N	1321 households observed on 8 periods.			
LM statistics for omission of $Y_{h,t-2}$ (critical value)	No		0.303 (5.99)	
Log Likelihood	-2749		-2378	

Note: Standard errors in parentheses, adjusted for clustering on households. *=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level. The Z vector of conditioning variables was reduced to ($Y_{h,0}$,$X_{h,1}$) due to weak within-group variance of the X variables. The controls that were insignificant in the preliminary regressions are dropped and help to identify the selection equation (region, household size, other labour force statuses). Wave dummies control for price variations.

There are two striking results. First, as expected, in specification **2** the coefficient on own lagged participation drops sharply, as compared to that in specification **1**, although it remains significant at all normal levels. The omitted individual fixed effect in specification **1** biases upwards the coefficient on lagged participation in the usual way.

Second, individual smoking participation is statistically independent of partner's smoking participation in specification **2**. The strong effects of partner's lagged smoking in specification **1** (a bivariate probit without individual fixed effects) entirely disappear in specification **2** where individual effects are modelled in a flexible manner.

These individual effects take two possible values per partner (-0.205 and -2.413 for women in column 4). Each couple has an estimated probability θ_{jk} of having the pair of individual effects (c_j, c_k). The average estimated probabilities for each of the four pairs is shown in the variable name column. We use this information to calculate an expected value of the individual effect for each respondent.[13] The estimated individual effects in columns 3 and 4 turn out to be positively correlated, with a correlation coefficient of 0.523.

The absence of correlation between partners' behaviours then reveals some deeper structure in the descriptive results presented in Section 3. Partners' behaviours are indeed correlated in the raw data, but only because their associated individual effects are not independent. These two results together are consistent with positive assortative matching on smoking, but do not support a decision-making model of health behaviour within couples.

Three additional points can be highlighted in Table 5. The positive significant Mills ratio in the first two columns shows that smoking and couple duration are correlated, when we do not control for individual effects. The Mills ratio becomes insignificant when we introduce individual random effects in columns 3 and 4, suggesting that the correlation between smoking and couple duration comes from the correlation of traits that are fixed over time (such as lifestyle choices). The presence of children has no significant effect on smoking participation. There is however a "pregnancy" effect on women's smoking. The fact that this only appears for women is consistent with an asymmetric gender effect of parental smoking on child health. Last, we note that specification **2** is not rejected by an LM test for the omission of two-period lagged participation. Other results in Table 5 show declining probability of smoking after age 50 for males. Also smoking is more prevalent amongst less-educated males, and lower amongst housewives.

6 Further results

The results in section 5 lead us to reject the second prediction, *i.e.* the bargaining explanation of spousal correlation in smoking, in favour of matching (our first prediction). This section discusses some alternative tests of the second and third predictions. We look at results for specific kinds of couples (new vs. old, and couples with and without children), and then at how individuals react to past health changes.

[13] For example, using the numbers in Table 5, a household with the average probability of belonging to each of the four classes would have an individual effect for the woman of (0.09+0.119)*-0.205 + (0.133+0.658)*-2.413 = -1.9515.

6.1 Couple heterogeneity

It can be argued that within-household interactions concerning lifestyle choices take place during the first years of the couple. Hence, the lifestyle choices of stable couples should already be well matched when we observe them in Table 5's regressions, so that we are unlikely to observe a decision-making effect. To test this argument, we augment Sample 2 with all couples from Sample 1 that are observed over two consecutive periods, but who are not necessarily stable over all the nine waves. This yields Sample 3, whose characteristics are described in Appendix A.[14] As Sample 3 is unbalanced, we use simple quit models (specification **3**) and individual probits with gaussian random effects (specification **4**), instead of the random effect bivariate probit model of Section 5.[15] The quit model can be thought of as a first-difference regression, eliminating fixed effects, as it is based on a change in smoking status. In these regressions, partner's past participation is interacted with an indicator for the duration of the cohabitation spell (at least three years vs. less than three years).

There is no difference in the estimated interaction effect between old and new couples in the quit equation. The estimates in specification **4** do not provide any significant evidence of stronger partner influence in new couples, although the estimated coefficients on partner's past participation are slightly higher. Note that male's past smoking has an effect on female's current smoking, but for old couples only. This result still holds when one estimates individual random effect probits on Sample 2, but vanishes when one accounts for the correlated effects as shown in Section 5.

---

[14] As expected, individuals in Sample 3 are younger, more likely to smoke, and less likely to be married.

[15] Individual probit models with gaussian random effects and control of the initial conditions problem yield results for Sample 2 that are qualitatively similar to those provided by the random effect bivariate probit model: prediction 2 is rejected (but at the 5% level only). We cannot estimate discrete bivariate fixed effect probits for the short duration couples as identification of this model requires transitions between smoking statuses.

Table 6. Old Couples & New Couples

Specification	**3: Quit probits**		**4: Gaussian Random-Effect Probits**	
Sample	*Sample 3*			
Equation	*Male*	*Female*	*Male*	*Female*
Past level of consumption	-0.031*** (0.004)	-0.044*** (0.005)	No	No
Past participation	No	No	1.483*** (0.087)	1.427*** (0.103)
Partner's past participation & cohabitation less than three years.	-0.183** (0.081)	-0.039 (0.087)	0.208 (0.171)	0.301 (0.187)
Partner's past participation & cohabitation more than three years.	-0.211*** (0.060)	-0.070 (0.064)	0.087 (0.135)	0.299** (0.138)
N	5115	4760	13908	13908

Note: Standard errors in parentheses, adjusted for clustering on households. *=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level. Other right-hand side variables as in Table 5.

Another couple characteristic which may change the way couples interact is the presence of children. In Section 2.3, we noted that the presence of children may reinforce the public good value of time spent with one's partner in good health.[16]

Table 7 below interacts lagged smoking statuses with dummies for the presence of children, in specification **2** of Table 5. The estimated correlations with partner's lagged smoking status remain insignificant.

Table 7. Spousal smoking correlation & children

Specification	**5: Discrete Random Effect Bivariate Probit**	
Sample	Sample 2	
Equation	*Male*	*Female*
Past participation & a child	1.959*** (0.146)	2.142*** (0.195)
Past participation & no child	1.575*** (0.321)	1.853*** (0.257)
Partner's past participation & partner has a child	0.041 (0.463)	0.051 (0.623)
Partner's past participation & partner doesn't have child	0.287 (0.479)	0.205 (0.556)

Note: Standard errors in parentheses, adjusted for clustering on households. *=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level. Other right-hand side variables as in Table 5.

[16] Using the marital and fertility history information in Wave 2 (including the number of adopted, step- and biological children), and information about the number of children at home, we know whether 98% of the individuals in sample 2 have already had a child at home (biological and adopted/step children are treated equally).

6.2 Health Changes

Our third prediction is that smoking status may be positively or negatively correlated with partner's past health changes, as a consequence of either correlated effects (social learning) or household decision-making. As such, the correlations of past health changes with smoking should be interpreted carefully. If both A and B smoke, a negative health shock for A will have an ambiguous effect on B's consumption: the expected future value of time spent with one's partner falls, which increases B's consumption (the decision-making argument), but B's subjective evaluation of the dangers of smoking rises (the social learning argument). An unanticipated negative health shock for a non-smoking A will however unambiguously increase B's consumption, since the social learning argument does not apply.

We report results from a discrete random effects bivariate probit on individual smoking status, as in Table 5, including health development variables, as measured by changes in subjective health status.[17] We denote all of *i*'s past health developments while smoking (both own and partner's) within a time period *j* (i.e. from *j-1* to *j*) by ΔH_{ij}. As in Clark and Etilé (2002), we explain current smoking status as a function of the sum of all past health changes: $\sum_{j=1}^{t-1} \Delta H_{ij}$. A positive correlation between these variables and smoking may be interpreted as learning about smoking's dangers.

Previous work using the same methods and dataset (Clark and Etilé, 2002) found some evidence that individual cigarette consumption reacts to own past health changes, but is largely independent of partner's health changes. Here we will reproduce this exercise, but with a far cruder binary measure of smoking. Table 8 shows the results.

Table 8: Smoking Equations with Health Changes.

Specification	**6: Discrete Random-Effect Bivariate Probit**	
Sample	Sample 2	
Equation	*Male*	*Female*
Past participation: $Y_{i,t-1}$	1.982*** (0.184)	2.031*** (0.323)
Partner's past participation: $Y_{-i,t-1}$	-0.020 (0.914)	0.184 (0.706)
Health status stayed good between *t-2* and *t-1*	Reference	Reference

[17] The subjective health status variable has five categories in the BHPS (excellent, good, fair, poor, very poor). We recoded excellent and good to "good", and fair, poor and very poor to "poor". At wave nine, the categories were somewhat different (excellent, very good, good, fair, poor), but the distribution of replies led us to keep the same grouping. We use this binary health variable (whereas Clark and Etilé, 2002, kept four categories) as our qualitative regressions do not allow efficient identification of a large number of health change dummies.

Health status changed from good to poor between *t-2* and *t-1*	-0.036 (0.084)	-0.009 (0.168)
Health status changed from poor to good between *t-2* and *t-1*	0.004 (0.049)	0.039 (0.110)
Health status stayed poor between *t-2* and *t-1*	-0.041 (0.043)	-0.066 (0.083)
Partner smoked at *t-1*, partner's health status stayed good between *t-2* and *t-1*	0.074 (0.098)	-0.018 (0.095)
Partner smoked at *t-1*, partner's health status changed from good to poor between *t-2* and *t-1*	-0.242 (0.307)	0.276 (0.331)
Partner smoked at *t-1*, partner's health status changed from poor to good between *t-2* and *t-1*	-0.063 (0.223)	-0.027 (0.443)
Partner smoked at *t-1*, partner's health status stayed poor between *t-2* and *t-1*	0.053 (0.222)	-0.135 (0.318)
Partner's health status stayed good between *t-2* and *t-1*	Reference	Reference
Partner's health status changed from good to poor between *t-2* and *t-1*	0.050 (0.327)	-0.072 (0.320)
Partner's health status changed from poor to good between *t-2* and *t-1*	0.081 (0.251)	0.073 (0.426)
Partner's health status stayed poor between *t-2* and *t-1*	-0.037 (0.221)	0.026 (0.289)
N	1321 households observed on 7 periods	

Note: Standard errors in parentheses, adjusted for clustering on households. *=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level. Other control variables as in Table 5.

We find no correlation between past health developments and current smoking status (controlling for own and partner's past participation). All of the estimated coefficients are insignificant in the discrete random-effect bivariate probit models. These regressions reveal little evidence of either social learning or bargaining.

This latter contrasts with the findings in Clark and Etilé (2002), where one's own health changes did matter. There are at least two potential reasons for this: Clark and Etilé consider four categories of health developments, interacted with age group and sex; and, perhaps most importantly, they model the level of daily cigarette consumption, not the participation decision in its own right. With respect to the last point, convex adjustment costs will render the level of consumption more malleable than the decision to smoke itself.

7 Conclusion

This paper has used nine waves of BHPS data to examine intra-spousal correlations in smoking behaviour. Perhaps this paper's most important contribution has been to the interpretation of observed correlations between spouses' behaviours. These can come about because partners' fixed traits are similar, as in matching models of marriage. Alternatively, household decision-making over health investments can lead couples to take similar smoking decisions. Last, individuals may learn about smoking's dangers by observing what happens to their smoking partner.

We first note that there is indeed a correlation in smoking status in the raw data, although there are some differences by sex. Further, a bivariate probit model, without controls for unobserved individual heterogeneity, reveals a positive correlation between partners' smoking participation: this is consistent with both matching and decision-making.

Our empirical approach has allowed us to distinguish between the three arguments above. Partners' smoking behaviours are statistically independent in a bivariate probit with individual random effects: all of the correlation in smoking status works through the correlation in individual effects. Further, we find very little evidence to support social learning in terms of smoking status. As such, we believe that the correlation in the raw data reflects matching on the marriage market, rather than household decision-making or learning within the couple. Smoking, and perhaps more generally health or lifestyle, seems to be one of the key domains over which sorting takes place in the marriage market, and the positive correlation shows that lifestyles are complements in the household production function.

There are at least two important consequences of our empirical analysis of interactions between household members. The first is purely descriptive or positive: if we are able to better identify the source of similarity in couples behaviours, we will build better, in the sense of more realistic and accurate, economic models of the household.

The second is normative: absent special cases, optimal policy is bound to depend on the nature of household interactions. In terms of the current paper's subject matter, we can ask whether it is sufficient, or more efficient, to target one person per household in terms of health education (or some other intervention), as opposed to all household members. Given the matching of partners' preferences for smoking, but only weak evidence of spillovers in cigarette consumption between partners during marriage, it seems essential to target both partners in order to reduce household smoking. Interventions targeting only the female partner (for instance during pregnancy) would not appear to be effective in reducing male smoking.

Appendix A: Descriptive Statistics

Sample	**1**		**2**		**3**	
Gender	*Males*	*Females*	*Males*	*Females*	*Males*	*Females*
N=	29692	33838	10568		19307	
Smokes at *t*	28.3%	26.7%	20.9%	20.3%	25.9%	24.1%
Quit between *t-1* and *t*	2.7%	2.3%	2.5%	2.5%	2.7%	2.7%
Age	44.7	46.4	49.6	47.3	47.6	45.2
Log individual yearly real income	9.07	8.40	9.39	8.35	9.31	8.34
Household Size = 2: ***reference***	34.9%	33.0%	39.9%	39.9%	42.8%	42.8%
Household Size = 3	20.1%	19.3%	20.8%	20.8%	21.0%	21.0%
Household Size = 4	20.9%	18.8%	26.7%	26.7%	24.0%	24.0%
Household Size = 5	8.3%	7.6%	10.3%	10.3%	9.4%	9.4%
Household Size = 6	2.8%	2.6%	2.3%	2.3%	2.8%	2.8%
Has at least one child	28.1%	31.0%	41.8%	41.9%	41.1%	41.1%
New child between *t-1* and *t*	3.5%	3.3%	3.4%	3.5%	4.6%	4.7%
Married: ***reference***	60.5%	53.1%	96.1%	96.1%	89.6%	89.6%
Living together	8.8%	7.7%	3.9%	3.9%	10.4%	10.4%
Not in couple	30.6%	39.3%	0.0%	0.0%	0.0%	0.0%
Manager (permanent)	13.9%	7.1%	18.9%	8.1%	17.0%	8.0%
Supervisor (permanent)	9.3%	7.5%	9.6%	8.7%	10.0%	8.8%
No responsibilities (permanent)	24.6%	25.4%	25.0%	31.6%	24.1%	29.7%
No responsibilities (temporary): ***reference***	7.4%	8.0%	5.6%	8.6%	6.7%	9.0%
Self-employed	11.5%	3.5%	13.4%	4.3%	13.7%	4.5%
Unemployed	5.7%	2.4%	3.2%	0.9%	4.2%	1.4%
Retired	18.2%	21.7%	20.1%	15.9%	18.7%	14.6%
Mother-at-home	0.5%	17.1%	0.4%	18.4%	0.4%	20.5%
School or training	4.1%	4.2%	0.3%	0.5%	0.5%	0.7%
Labour force status: not defined	4.8%	3.2%	3.7%	2.9%	4.6%	2.8%
Education under A-level	33.0%	40.5%	33.7%	39.5%	33.1%	38.2%
Education A-level or over	67.0%	59.5%	66.3%	60.5%	66.9%	61.8%
Subjective health status = "good" or "excellent"	69.7%	65.0%	73.2%	70.0%	70.2%	67.6%

Appendix B – Additional results

Table B1. Instrumental regression for selection bias.

Selection	**Sample 1 into Sample 2**	
Equation	*Male*	*Female*
Unemployment rate	0.008 (0.013)	-0.015 (0.014)
Partner unemployment rate	-0.005 (0.015)	0.011 (0.012)
Manager (permanent)	0.638*** (0.048)	0.054 (0.044)
Supervisor (permanent)	0.407*** (0.049)	0.029 (0.043)
No responsibilities (permanent)	0.337*** (0.045)	0.086** (0.038)
Self-employed	0.562*** (0.048)	0.174*** (0.051)
Unemployed	-0.090* (0.054)	-0.729*** (0.063)
Retired	0.418*** (0.047)	-0.412*** (0.039)
Mother-at-home	0.107 (0.118)	0.015 (0.039)
School or training	-1.320*** (0.083)	-1.458*** (0.067)
Labour force status: not defined	0.325*** (0.055)	-0.219*** (0.054)
Education = A-level	-0.094*** (0.020)	0.024 (0.018)
Education>A – level	0.080*** (0.019)	0.129*** (0.020)
Controls for years and regions	Yes	Yes
N	29692	33838
LnL (LnL_0)	-19249 (-20333)	-21282 (-22547)

Note: *=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level, N.S.=insignificant at the 10% level. LnL_0: log-likelihood for the constant-only model.

Table B2. Multinomial logit for the distribution of the individual effects.

Point of support No.	1,2		2,1		2,2	
Equation	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
$Y_{h,O}$	1.052 (0.860)	-6.523*** (0.983)	-5.155*** (0.749)	0.523 (0.926)	-4.329*** (0.756)	-5.394*** (0.817)
Mills ratio	-0.055 (1.468)	1.890 (1.401)	-0.054 (1.724)	-0.193 (1.434)	-2.128 (1.647)	0.514 (1.406)
Initial income	-0.435 (0.521)	0.719 (0.356)	-0.366 (0.479)	0.432 (0.368)	0.124 (0.494)	0.775** (0.337)
Initial age/10	0.580 (3.476)	-1.338 (3.257)	2.529 (3.547)	-2.767 (3.456)	2.018 (3.290)	-2.296 (3.263)
Initial $age^2/100$	-0.117 (0.354)	0.185 (0.358)	-0.216 (0.360)	0.294 (0.364)	-0.160 (0.327)	0.230 (0.347)
Initial education ≥ A-level	0.208 (0.682)	0.367 (0.707)	0.045 (0.662)	0.561 (0.699)	0.235 (0.610)	0.495 (0.661)
Newborn child between *t-1* and *t*	0.726 (1.057)		0.999 (1.662)		1.237 (1.335)	
Has at least one child at home	0.013 (0.749)		0.622 (0.681)		-0.212 (0.664)	
Initial marital status: live together	-1.859 (1.225)		0.067 (1.345)		-0.021 (1.002)	
Initial household size=5	-1.004 (0.854)		-1.334* (0.793)		-1.143 (0.826)	
Constant	0.670 (6.684)		0.758 (6.216)		-0.571 (6.276)	

Note: Point number 1,1 (*cf.* Table 5 above) is the reference for the estimation of the multinomial logit probabilities conditional on $(Y_{h,0},X_{h,1})$ (see equation **(4)**).

References

Becker, G.S., 1974. A Theory of Marriage: Part I. Journal of Political Economy 81, 813-846.

Becker, G.S., Murphy, K.M., 1988. A Theory of Rational Addiction, Journal of Political Economy 96, 675-700.

Bergstrom, T., 1997. A Survey of Theories of the Family. In: Rosenzweig, M.R., Stark, O. (Eds.), Handbook of Population and Family Economics Volume 1A. Elsevier Science B.V., Amsterdam, pp. 21-74.

Bolin, K., Jacobson, L., Lindgren, B., 2001. The family as the health producer – when spouses are Nash-bargainers. Journal of Health Economics 20, 349-362.

Bolin, K., Jacobson, L., Lindgren, B., 2002. The family as the health producer – when spouses act strategically. Journal of Health Economics 21, 475-495.

Celeux, G., Chauveau, D., Diebolt, J., 1995. On Stochastic Versions of the EM algorithm. Institut National de Recherche en Informatique et en Automatique, Working Paper No. 2514.

Clark, A.E., Etilé, F., 2002. Do Health Changes Affect Smoking? Evidence from British Panel Data. Journal of Health Economics 21, 533-562.

Clark, A.E., Lohéac, Y., 2005. It wasn't me, It was them! A Study of Social Influence in Risky Behaviour by Adolescents. IZA, Discussion Paper No. 1573.

Currie, J., Madrian, B., 1999. Health, Health Insurance and the Labour Market. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics Volume 3C. Elsevier Science B.V., Amsterdam, pp. 3309-3416.

Contoyannis, P., Jones, A., 2004. Socio-economic status, health and lifestyle. Journal of Health Economics 23, 965-995.

Dempster, A.P., Laird N.M., Rubin, P.B., 1977. Maximum-Likelihood from incomplete data via the E-M algorithm. Journal of the Royal Statistical Society B 39, 1-38.

Deb, P., Trivedi P.K., 1997. Demand for Medical Care by the Elderly: A Finite Mixture Approach. Journal of Applied Econometrics 12, 313-336.

Farrell, L., Shields, M., 2002. Investigating the economic and demographic determinants of sporting participation in England. Journal of the Royal Statistical Society A 165, 335-348.

Fernández, R., Guner, N., Knowles, J. 2005. Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality. Quarterly Journal of Economics 120, 273-344.

Jedidi, K., Jagpal, H.S., DeSarbo, W.S., 1997. Finite-Mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity. Marketing Science 16, 39-59.

Jenkins, S., Osberg, L., 2005. Nobody to play with? The implications of leisure coordination. In: Hamermesh, D., Pfann, G. (Eds.), The Economics of Time Use. Elsevier Science B.V., Amsterdam, pp. 113-145.

Kan, M.-Y., Heath, A., 2003. The political attitudes and choices of husbands and wives. University of Oxford, mimeo.

Keane, M., 1992. A Note on Identification in the Multinomial Probit Model. Journal of Business and Economic Statistics 10, 193-200.

Latkin, C., Mandell, W., Oziemkowska, M., Celentano, D., Vlahov, D., Ensminger, M., Knowlton, A., 1995. Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. Drug and Alcohol Dependence 38, 1-9.

Leonard, K., Mudar, P., 2003. Peer and Partner Drinking and the Transition to Marriage: A Longitudinal Examination of Selection and Influence Processes. Psychology of Addictive Behaviors 17, 115-125.

Levine, P.B., Gustafson, T.A., Velenchik, A.D., 1997. More Bad News for Smokers? The Effects of Cigarette Smoking on Wages. Industrial and Labor Relations Review 50, 493-509.

Manser, M., Brown, M, 1980. Marriage and household decision-making: a bargaining analysis. International Economic Review 21, 31-44.

McElroy, M., Horney, M., 1981. Nash-bargained decisions: toward a generalization of the theory of demand. International Economic Review 22, 333-349.

Manski, C., 1995. Identification Problems in the Social Sciences. Harvard University Press, Cambridge, MA..

Rose, E., 2004. A Joint Econometric Model of Marriage and Partner Choice. University of Washington, mimeo.

Van Ours, J., 2004. A Pint a Day Raises a Man's Pay; But Smoking Blows that Gain Away. Journal of Health Economics 23, 863-886.

Vella, F., 1993. A Simple Estimator for Simultaneous Models with Censored Endogenous Regressors. International Economic Review 34, 441-457.

Wedel, M. , DeSarbo, W.S., Bult, J.R., Ramaswamy, V., 1993. A Latent Class Poisson Regression Model for Heterogeneous Count Data. Journal of Applied Econometrics 8, 397-411.

Weiss, Y., 1997. The Formation and Dissolution of Families: Why Marry? Who Marries whom? And What Happens Upon Divorce. In: Rosenzweig, M.R., Stark, O. (Eds.), Handbook of Population and Family Economics Volume 1A. Elsevier Science B.V., Amsterdam, pp. 81-124.

Wilson, S., 2002. The Health Capital of Families: An Investigation of the Inter-Spousal Correlation in Health Status. Social Science and Medicine 55, 1157-1172.

Wooldridge, J.M., 2002a. Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge, MA.

Wooldridge, J.M., 2002b. Simple solutions to the initial conditions problem in dynamic non-linear panel data models with unobserved heterogeneity. IFS/UCL CEMFI Working Paper 18/02.